

# ***AIDS Public Information Data Set***

---

**Data through December 2000**

**U.S. Department of Health and Human Services**

Public Health Service  
Centers for Disease Control and Prevention  
National Center of HIV, STD, and TB Prevention  
Division of HIV/AIDS Prevention



# About this Data Set

The *AIDS Public Information Data Set* is computer software designed to run on an Microsoft Windows microcomputer, and contains information abstracted from acquired immunodeficiency syndrome (AIDS) cases reported in the United States. The data set is created each year by the Division of HIV/AIDS Prevention, National Center for HIV, STD, and TB Prevention, Centers for Disease Control and Prevention (CDC) and contains information extracted from CDC's national AIDS surveillance data base. Suggested citation: *Centers for Disease Control and Prevention. AIDS Public Information Data Set, December, 2000.*

In December 1995, the software was modified to add data from all metropolitan areas with 500,000 or more population, for metropolitan areas with 100,000 or more population from selected states, and for individual counties or health districts from selected states. To protect the confidentiality of the data, some information was removed from the data set. Month of death, survival time, central versus outlying portion of metropolitan areas, reporting delay adjustments for death dates, and information on individual AIDS-indicator conditions were removed from the data set. The December 1996 edition added information on patient vital status. The December 1998 edition added percentage calculation to each tabulation. Additional information is contained in the on-line help files.

The *AIDS Public Information Data Set* contains data in two formats. The first format consists of a rectangular data file of 16 variables extracted from CDC's national AIDS data base. One-way and two-way cross tabulations of any of these variables can be displayed on your computer screen. The second format consists of a series of state, metropolitan statistical area (MSA), and county/health district tables, containing information on 8 variables included in the rectangular data file plus a location variable. There is one set of tables for the entire United States, one set for each state, one set for each MSA, and one set for each county/health district. The rectangular data file, without the state or MSA tables, is also available as an ASCII data file.

To request a copy of this data set, contact the Statistics and Data Management Branch, Division of HIV/AIDS Prevention, Mailstop E-48, Centers for Disease Control and Prevention, Atlanta, GA, 30333, telephone (404) 639-2020. You can also download the software from the Internet by linking to <http://www.cdc.gov/hiv/software.htm>.

This manual describes the data set. It is divided into three sections and three appendices. On-line help screens provide additional information.

Section 1, *AIDS Surveillance in the United States*, describes the data collection process and the effect changes in this process may have on data analysis and interpretation. The section reviews the source of AIDS surveillance data and describes which patients are included in the CDC definition. It also discusses reporting delays and reporting completeness.

Section 2, *Data File Variables and Coding Schemes*, lists the variables included on the rectangular data file and describes each variable's coding scheme.

Section 3, *State, MSA, and County Tables*, describes the variables included on the state, MSA, and county/health district tables.

*Appendix A: Installation*, describes how to load and run this program on your computer. It also suggests computer hardware and software you can use to analyze the data.

*Appendix B: Metropolitan Statistical Areas* lists the MSAs included in the data set.

*Appendix C: Health Districts* lists the counties which comprise each health district included in the data set.

# Assurance of Confidentiality

The data files on the enclosed CD contain information abstracted from acquired immunodeficiency syndrome (AIDS) case reports received from state and local health departments, who voluntarily report cases of AIDS to CDC. Case reports do not include patient or physician names or other personal identifiers. The data are protected under the Assurance of Confidentiality (Sections 306 and 308(d) of the Public Health Service Act, 42 U.S.C. 242k and 242m(d)), which prohibits disclosure of any information that could be used to directly or indirectly identify patients. The statistical data contained in the *AIDS Public Information Data Set* are being released for public use in accordance with the assurance and do not identify patients directly, nor do they contain information that can identify patients indirectly.



# AIDS Surveillance in the United States

## Background

In 1981, after early reports of *Pneumocystis carinii* pneumonia, Kaposi's sarcoma, and other opportunistic infections in young homosexual men in Los Angeles, New York, and San Francisco, the Centers for Disease Control and Prevention (CDC) began surveillance for a newly recognized constellation of diseases, now termed the acquired immunodeficiency syndrome (AIDS). CDC developed a surveillance case definition for this syndrome and initially received case reports directly from health care providers and state and local health departments. As the epidemic spread, state and local health departments assumed responsibility for AIDS surveillance, and by 1985 all states had regulations requiring physicians and other health care providers to report AIDS cases directly to the state or local health department. These health departments then share the reports with CDC, which produces the national AIDS surveillance data set.

The goals of AIDS surveillance have been to monitor both trends in AIDS cases and the scope of severe morbidity due to infection with the human immunodeficiency virus (HIV). AIDS surveillance data are used to allocate resources for patient care, target HIV prevention programs, and evaluate the impact of public health recommendations. Advances in the understanding of the epidemiology and manifestations of HIV infection and changing diagnostic practices, however, present multiple challenges to those analyzing and interpreting the AIDS surveillance data. The following are a few examples:

- ! A wide variety of persons are at risk for HIV, including men who have sex with men, injecting drug users, person who received a transfusion or who were tissue transplant recipients before March 1985, heterosexual partners of infected persons, children born to infected mothers, and persons with mucous membrane or percutaneous exposure to blood or body fluids of infected persons (e.g., health care workers). Because men who have sex with men comprise such a large proportion of the total number of AIDS cases, trends in this subgroup will overshadow those in other groups unless the data are examined separately. Analysis of data, without regard to specific subgroups, may conceal information or lead to misinterpretation of the data.
- ! The etiologic agent of AIDS, HIV, has been identified, and diagnostic tests for infection with this virus have been developed. As a result, the surveillance of AIDS, initially dependent on the presence of certain indicator diseases specific for the infection, was expanded in 1985, 1987, and 1993 to include additional conditions (some conditions may be less specific for HIV infection) in the presence of laboratory evidence for infection, and in 1993 to include HIV-infected persons with laboratory evidence of severe immunosuppression. The addition of these conditions to the AIDS case definition has affected trends in reported AIDS cases, as well as trends in reporting of AIDS-defining opportunistic conditions.

- ! Diagnostic practices have changed over time and vary geographically. AIDS is now a common diagnosis in many hospitals and clinics, and definitive diagnostic tests for manifestations of HIV infection (e.g., *Pneumocystis carinii* pneumonia or esophageal candidiasis) may not be done. HIV testing is not available for all patients and some patients choose not to be tested. Geographic variations in diagnostic practices and surveillance procedures, and changes over time could markedly affect trends in AIDS surveillance.

## Source of AIDS Surveillance Data

CDC maintains national AIDS surveillance through receipt of AIDS case reports submitted by individual state and local health departments. Health departments report cases electronically through a CDC-developed microcomputer system. All 50 states, the District of Columbia, U.S. dependencies and possessions, and independent nations in free association with the United States (Puerto Rico, the U.S. Virgin Islands, Guam, American Samoa, the Republic of Palau, the Republic of the Marshall Islands, the Commonwealth of the Northern Mariana Islands, and the Federated States of Micronesia) report AIDS cases to CDC.

Although state and local health departments share AIDS surveillance data with CDC, the responsibility and authority for AIDS surveillance rests with the individual health departments. Like any reportable disease, the completeness of AIDS reporting reflects how actively health departments solicit case reports. Historically, disease surveillance systems have been categorized as passive or active, i.e., health departments may passively receive case reports from health care providers, depending on health care providers to know and comply with reporting requirements; or they may actively contact and interact with health care facilities or individual providers to stimulate disease reporting, sometimes directly assuming the primary responsibility of reporting cases from large or high-volume institutions.

CDC provides funding and technical assistance to health departments to actively stimulate AIDS case reporting and has encouraged them to take an active rather than passive approach to AIDS surveillance. Through surveillance cooperative agreements supported by CDC, health departments are encouraged to identify health care facilities that serve AIDS patients and work closely with these facilities to encourage reporting. They are also encouraged to send newsletters to health care providers and attend professional organization meetings, and to use other data sources to identify AIDS cases, including death certificates, laboratory reports, and tuberculosis and tumor registries. States vary in the structure and organization of their surveillance systems and, therefore, in the completeness of their case reporting (see below).



## Case Definition

Before HIV was identified as the etiologic agent for AIDS, CDC defined a case of AIDS (for surveillance purposes) as a disease, at least moderately indicative of a defect in cell-mediated immunity, occurring in a person with no known cause for diminished resistance to the disease. Such diseases included *Pneumocystis carinii* pneumonia, Kaposi's sarcoma, and many other serious opportunistic infections (see *American Journal of Medicine*, March 1984, pages 493-500). With identification of HIV as the causative agent for AIDS and the availability of laboratory tests to detect HIV antibody, the case definition was expanded to reflect an increased understanding of HIV infection in 1985 (see CDC's *Morbidity and Mortality Weekly Report*, June 28, 1985, pages 373-375) and in 1987 (see *Morbidity and Mortality Weekly Report*, August 14, 1987, supplement, pages 3S-15S). These revisions applied to persons with laboratory evidence for HIV infection. Among diseases added in 1985 were disseminated histoplasmosis, chronic isosporiasis, and certain non-Hodgkin's lymphomas. Among those added in 1987 were extrapulmonary tuberculosis, HIV encephalopathy, and HIV wasting syndrome. In children, recurrent, serious bacterial infections were also added. In addition, the 1987 revision allowed certain indicator diseases to be diagnosed presumptively based on clinical presentation rather than "confirmed" by laboratory or diagnostic methods.

To be consistent with standards of medical care for HIV-infected persons and to more accurately reflect the number of persons with severe HIV-related immunosuppression who are at highest risk for HIV-related morbidity and most in need of close medical follow-up, the surveillance definition was expanded on January 1, 1993 (see CDC's *Morbidity and Mortality Weekly Report, Recommendations and Reports*, December 18, 1992). This expansion includes all HIV-infected adults and adolescents who have less than 200 CD4<sup>+</sup> T-lymphocytes/ $\mu$ L or a CD4<sup>+</sup> T-lymphocyte percent of total lymphocytes less than 14, or who have been diagnosed with pulmonary tuberculosis, invasive cervical cancer, or recurrent pneumonia. The addition of pulmonary tuberculosis, recurrent pneumonia, and invasive cervical cancer in HIV-infected adults and adolescents to the 23 clinical conditions listed in the 1987 surveillance definition reflects their documented or potential importance in the HIV epidemic.

While the reported incidence of AIDS increased only 3 to 4 percent as a result of the 1985 revision, the 1987 revision greatly increased the numbers of reported cases. Roughly one fourth of all adults/adolescents who were both diagnosed and reported in the year following the 1987 revision were reported based only on the additional criteria included in the 1987 revision. Furthermore, the proportion of cases meeting only the revised criteria was higher in Hispanics and non-Hispanic blacks than in non-Hispanic whites, higher in heterosexual injecting drug users, and lower in men who have sex with men. The 1993 revision has had substantial impact on the number of reported cases. The immediate increase in case reporting was largely attributed to the addition of severe immunosuppression to the definition; a smaller impact was due to the addition of pulmonary tuberculosis, recurrent pneumonia, and invasive cervical cancer, since many persons with these diseases also have a CD4<sup>+</sup> T-lymphocyte count of less than 200 cells/ $\mu$ L. The early effects of expanded surveillance were greater than long-term effects because prevalent as well as incident cases of immunosuppression were reported after implementation of the expanded surveillance case definition. In

recent years, the effect on the number of reported cases has been smaller. Due to the large number of cases reported based on criteria in only the revised case definitions and to the inconsistent use of the revised case definitions in different populations, analyses of trends in AIDS cases must take these revisions into account.

## Case report form

Separate case report forms are used for pediatric patients (patients less than 13 years of age at the time of diagnosis) and adult/adolescent patients (patients 13 years of age or older at the time of diagnosis). Although the forms are similar, the pediatric form includes behavioral risk information on the child's mother. These forms are completed by the health care provider or by the AIDS surveillance staff in the local or state health department. In addition, a laboratory report of an AIDS-defining condition sent to health departments may initiate a case report. In these cases, follow-up with the health care provider is required to obtain complete information.

Names are retained by the state or local health department and are converted to an alpha-numeric code called "soundex" for use by CDC. CDC does not receive names of persons with AIDS. Because more than one state may report an individual case, CDC screens reported cases by soundex code, date of birth, sex, and state of residence to cull presumed duplicate reports. States also cooperate in this process by reporting out-of-jurisdiction cases to the patient's state of residence.

The variables available on the AIDS data set are listed in the next section. However, a few deserve special comment.

- ! *Vital status.* Patients survive for a variable amount of time following the diagnosis of AIDS. Because death usually occurs after the initial report to CDC, case reports may not be updated to reflect the change in vital status. As a result, reporting of deaths among AIDS patients may be delayed or incomplete. However, states are required to perform periodic reviews of death certificates and state death registries to identify unreported cases, and to update vital status of known cases. In addition, 16 states participated in a special project to match their case registries to the National Death Index to assess the completeness of reporting and to identify deaths among cases that died out-of-jurisdiction.
- ! *Exposure category.* Some patients may have more than one mode of exposure to HIV. For surveillance purposes, AIDS cases are counted only once in a hierarchy of exposure categories. Persons with more than one reported mode of exposure are listed in the category that appears first in the exposure hierarchy, except for men with both a history of sexual contact with other men and injecting drug use. They make up a separate exposure category.
- ! *AIDS definition category.* Patients may develop additional conditions indicative of AIDS after

their initial AIDS diagnosis. The case report form may not be updated to reflect additional conditions. Some persons reported as meeting only the immunologic criteria may have concurrent or prior opportunistic infections or conditions that are not included in the case report. Therefore, cases reported as meeting only the criteria added to the case definition in 1993 may include persons who meet the criteria in 1987 definition.

- ! *Date of diagnosis.* CDC collects dates of diagnosis for each AIDS-indicator disease, and, for patients with severe immunosuppression, the date of the CD4<sup>+</sup> T-lymphocyte test. From this information, a single date of diagnosis is calculated for each patient; it is the earliest of these dates.

## Delay in Reporting

The timeliness of AIDS case reporting to CDC is dependent on a number of factors, including the volume of cases reported from a state or locality, the cooperation of health care providers and medical institutions, the availability of staff to complete case report forms, and changes in the case definition. In many instances initial case reports are incomplete and require additional follow-up by state and local health department staff, including reviews of other record systems and contact with health care providers.

Based on estimates calculated using AIDS surveillance data reported between 1995 and 2000, about 50 percent of all cases were reported to CDC within 4 months of the date of diagnosis, but about 20 percent were reported more than 1 year after diagnosis. Delays vary widely among geographic, age, exposure, sex, and racial/ethnic categories. They are substantially longer for pediatric cases and shorter for AIDS cases previously reported with HIV infection, for example. Due to the reporting delay, the number of cases diagnosed during any period often exceeds the number reported during that period. This is particularly important in examining trends over time, since many cases in recent periods of time will not yet be reported.

To account for delays in the reporting of cases, the variable *adjwgt* is included in the data set. This variable may be used to weight each case on the data set and obtain adjusted case counts. For example, summing *adjwgt* for cases would estimate the number of cases diagnosed through the time period covered by the data set that will eventually be reported to CDC. To use this variable, select the adjustment weight option from the *Tools* menu. Once you turn the option on, all subsequent tabulations will be adjusted for reporting delay. The adjustment weight and resulting tabulations are not reliable for cases diagnosed during the most recent 6 to 9 months.

## Effect of CD4 Reporting on AIDS Case Trends

As a result of the case definition change in 1993, trends in AIDS case counts showed an artifactual peak early in 1993, even after adjustment for reporting delay. To examine trends over time using

a constant case definition, i.e., diagnoses of opportunistic illnesses that were included in the 1987 or the 1993 case definition, CDC developed methods that estimated incidence of 1987 or 1993 definition opportunistic infections for cases that met only the 1993 immunologic (CD4<sup>+</sup>) criteria. These estimates showed that the number of diagnoses of AIDS-defining opportunistic infections increased during 1992 and 1993 by approximately 2 percent and 3 percent, respectively (see *Morbidity and Mortality Weekly Report*, November 18, 1994). The temporary distortion of the AIDS incidence curve caused by the 1993 expansion of the AIDS case definition had almost entirely waned by 1996.

## Effect of Therapy on AIDS Incidence

Continuing the pattern first observed from 1995 to 1996, AIDS incidence decreased again from 1996 to 1997 and from 1997 to 1998. These decreases are mostly due to the effect of therapies for HIV infection and AIDS, which have altered the natural history of HIV infection and slowed progression to AIDS. AIDS incidence increasingly represents persons who were not diagnosed with HIV infection until they developed AIDS, persons who did not access treatment, or persons for whom treatment failed. Caution should be used when interpreting trends in AIDS incidence; the contribution of these effects to the AIDS incidence curve is currently being evaluated. See *Morbidity and Mortality Weekly Report*, September 19, 1997 and April 24, 1998.

## Early Reporting Dates

Before 1990, CDC occasionally received reports on patients before they met the CDC AIDS case definition. If such patients were later diagnosed with AIDS, the diagnosis date on their record (when they first met the CDC definition) would be after the report date (when CDC first received information about the patient). Such records should be excluded from certain analyses, such as survival analysis and analysis of reporting delay. CDC's AIDS surveillance data base no longer receives reports on patients who do not meet the AIDS case definition.

## Follow-up of Reported AIDS Cases

AIDS case records maintained at CDC contain all information reported to date from state and local health departments. As patients progress through their illness, additional conditions may be reported, or the patient's vital status may change. However, not all health departments have the resources to routinely follow-up patients for additional information. For this reason and because many patients move out of the reporting health department's jurisdiction, CDC records do not always contain all current information for each patient.

AIDS cases reports that do not include mode of HIV exposure information are routinely followed up by state and local health departments. As of December 1999, excluding cases which were not yet investigated, mode of exposure information has been identified for 78 percent of cases. Twenty-one percent of cases were closed with incomplete information because the patient died, declined interview, or was lost to follow-up; 1 percent of cases remained without a reported risk for HIV infection after complete investigation (see Centers for Disease Control and Prevention. *HIV/AIDS Surveillance Report*, 1999;11(no.2):27). The demographic profile of persons who remain without risk information is more similar to that of other persons reported with AIDS than with the general U.S. population.

## Evaluation of AIDS Surveillance

Cases of AIDS may not be reported to CDC for a variety of reasons. The diagnostic tests needed to confirm the diagnosis of certain AIDS-indicator conditions may not be performed, or physicians and hospital personnel may fail to report cases to the health department. Further, some patients with HIV disease may be ill or die from diseases or conditions not included in the current AIDS surveillance definition or from causes unrelated to their HIV infection.

Both CDC and state and local health departments have commissioned a variety of studies to evaluate the completeness of AIDS surveillance. Most evaluation projects have used alternate data resources if they are independent of routine case finding, such as death certificates, hospital discharge records, and laboratory records. Individual records from these alternate sources have then been matched against records in AIDS surveillance data bases. If an alternative source is found to be a productive source of case reports, it may be added to routine case finding methods. Evaluation projects have varied in size and scope (e.g., varying numbers of ICD-9 codes from death certificates or computerized discharge records), geographic area covered, detection of both inpatient and outpatient cases, and time frames. In general, evaluation studies suggest that reporting of AIDS cases is fairly complete; but, depending on the setting and evaluation method used, the level of reporting completeness may vary. High prevalence areas for AIDS appear to have more complete reporting than low prevalence areas. Following implementation of active case finding under the 1987 case definition, with funding support from CDC, completeness of case reporting increased in most areas and was estimated to be more than 85 percent complete (see *Journal of Acquired Immunodeficiency Syndrome*, 1992;5:257-64 and *American Journal of Public Health* 1992;82:1495-99).

## Summary

Public health surveillance represents an ongoing and regular collection, analysis, interpretation, and application of health data for disease prevention and control. AIDS surveillance, like other national surveillance efforts, depends on health care providers and the state and local health departments and, thus,

requires a balance between information needs versus practical limitations. AIDS surveillance in the United States represents an unprecedented public health enterprise and has achieved an unusually high degree of completeness. In addition, surveillance has changed as understanding of AIDS and HIV infection have grown. Users of the public information data set should be familiar with the characteristics of public health surveillance in general as well as with the evolution of AIDS surveillance.

# Data File Variables and Coding Schemes

The rectangular data file included in the *AIDS Public Information Data Set* contains one line of data for each AIDS case reported to CDC. Each line contains 35 columns. The columns contain 16 variables extracted from CDC's national AIDS data set.

Column	Variable	Description
1	<i>age</i>	Age group at diagnosis of the first AIDS-indicator opportunistic condition
2	<i>sexclass</i>	Sexual classification of patient
3	<i>race</i>	Race of patient
4	<i>categ</i>	Indicates which of the CDC AIDS case revisions the patient meets
5-10	<i>dxdate</i>	Month of diagnosis of first AIDS-indicator opportunistic condition
11-16	<i>repdate</i>	Date when CDC first received information about the case
17	<i>death</i>	Vital status of patient
18-19	<i>exposure</i>	Mode of exposure to HIV
20	<i>multrisk</i>	Indicates if patient had more than one risk of exposure to HIV
21	<i>birth</i>	Country of birth
22	<i>sexbi</i>	Sex with a bisexual man (women only)
23	<i>sexiv</i>	Sex with an injecting drug user
24	<i>sexother</i>	Sex with a person with hemophilia or with a transfusion recipient
25	<i>sexhiv</i>	Sex with a person known to be infected with HIV or to have AIDS, but whose mode of exposure is unknown
26-31	<i>adjwgt</i>	Reporting delay adjustment weight
32-35	<i>msa</i>	Region of residence at diagnosis of AIDS

Each of these variables is coded alpha-numerically. The codes used in the *AIDS Public Information Data Set* are described below.

## Age (column 1)

This variable contains the patient's age when he or she was first diagnosed with an AIDS-indicator disease.

0 = Less than 1 year old

1 = 1 to 12 years old

2 = 13 to 19 years old

3 = 20 to 24 years old

4 = 25 to 29 years old

5 = 30 to 34 years old

6 = 35 to 39 years old  
or age is missing  
7 = 40 to 44 years old  
8 = 45 to 49 years old  
9 = 50 to 54 years old  
A = 55 to 59 years old  
B = 60 to 64 years old  
C = 65 years old or older

### Sexclass (column 2)

Adult/adolescent males are classified according to their sexual orientation.

1 = Adult/adolescent male who has sex only with other men or sex is missing, or sexual orientation is missing  
2 = Adult/adolescent male who has sex with both men and women  
3 = Adult/adolescent heterosexual male or pediatric male  
4 = Female (both adult/adolescent and pediatric)

### Race (column 3)

1 = White (not Hispanic)  
2 = Black (not Hispanic)  
3 = Hispanic  
4 = Asian/Pacific Islander  
5 = American Indian/Alaskan Native  
9 = Unknown

### Categ (column 4)

This variable reflects changes made over time to the CDC surveillance definition for AIDS. Only cases meeting the current (1993) surveillance definition are included in this data set. *Categ* indicates whether the patient also met the pre-1985, 1985, or 1987 surveillance definition, and whether the diagnosis, if it meets the 1987 or 1993 definition, was definitive or presumptive. Cases that meet more than one of these surveillance definitions are classified into the category listed first. For more information about the 1993 definition, see *Morbidity and Mortality Weekly Report*, December 18, 1992, Recommendations and Reports.



- 1 = Case meets the pre-1985 surveillance definition
- 2 = Case meets the 1985 surveillance definition
- 3 = Case meets the 1987 surveillance definition and was diagnosed definitively
- 4 = Case meets the 1987 surveillance definition and was diagnosed presumptively
- 5 = Case meets the 1993 surveillance definition: pulmonary tuberculosis, recurrent pneumonia, and/or cervical cancer (definitive diagnosis)
- 6 = Case meets the 1993 surveillance definition: pulmonary tuberculosis and/or recurrent pneumonia (presumptive diagnosis)
- 7 = Case meets the 1993 surveillance definition, severe HIV-related immunosuppression

#### Dxdate (columns 5 through 10)

This variable contains the year and month in which the first AIDS-indicator condition was diagnosed. Columns 5 through 8 contain the year; columns 9 and 10 contain the month. Cases diagnosed before 1982 are coded as “198199.” Cases whose month of diagnosis is unknown are coded as “99” in the month portion of this variable.

#### Reptime (columns 11 through 16)

This variable contains the year and month in which CDC received the case report. Columns 11 through 14 contain the year; columns 15 and 16 contain the month. Cases reported during 1981 are coded as “198199.”

#### Death (column 17)

- 0 = CDC has not received a death notification for this case
- 1 = CDC has been notified that this patient died

Patients diagnosed during the 2 most recent years are coded as “0” regardless of the patient's vital status. AIDS prevalence rates calculated for the most recent two-year period should be interpreted with caution. The rates calculated will be artificially high because all persons diagnosed in this period are coded with a vital status of “0” (alive), even if a death has been reported to CDC for that person. This is to prevent inadvertent indirect identification of any record by linking a death date inferred from this data set to other publically available data sets which contain death dates on individuals. For more information on trends in AIDS, see *Morbidity and Mortality Weekly Report*, September 19, 1997 and April 24, 1998.

## Exposure (columns 18 and 19)

For surveillance purposes, AIDS cases are counted only once in a hierarchy of exposure categories. Persons with more than one reported mode of exposure to HIV are classified in the exposure category listed first in the hierarchy, except for men with both a history of sexual contact with other men and injecting drug use. They make up a separate exposure category. Persons with multiple reported modes of exposure are indicated in the variable *multirisk*.

“Men who have sex with men” cases include men who report sexual contact with other men (i.e., homosexual contact) and men who report sexual contact with both men and women (i.e., bisexual contact). “Heterosexual contact” cases are in persons who report specific heterosexual contact with a person with, or at increased risk for, HIV infection (e.g., an injecting drug user).

Adults/adolescents born, or who had sex with someone born, in a country where heterosexual transmission was believed to be the predominant mode of HIV transmission (formerly classified as Pattern-II countries by the World Health Organization) are no longer classified as having heterosexually acquired AIDS. Similar to case reports for other persons who are reported without behavioral or transfusion risks for HIV, these reports are now classified (in the absence of other risk information which would classify them into another exposure category) as “no risk reported or identified” (see *Morbidity and Mortality Weekly Report*, March 11, 1994). Children whose mother was born, or whose mother had sex with someone born, in a Pattern-II country are now classified (in the absence of other risk information which would classify them into another exposure category) as “Mother with/at risk for HIV infection: has HIV infection, risk not specified.”

“Risk not reported or identified” cases are in persons with no reported history of exposure to HIV through any of the routes listed in the hierarchy of exposure categories. Risk not reported or identified cases include persons who are currently under investigation by local health department officials; persons whose exposure history is incomplete because they died, declined to be interviewed, or were lost to follow-up; and persons who were interviewed or for whom other follow-up information was available and no exposure mode was identified. Persons who have an exposure mode identified at the time of follow-up are reclassified into the appropriate exposure category.

### Adult/adolescent exposure categories

- 1 = Men who have sex with men
- 2 = Injecting drug use
- 3 = Men who have sex with men and inject drugs

- 4 = Hemophilia/coagulation disorder
- 5 = Heterosexual contact with a person with, or at increases risk for, HIV infection
- 7 = Receipt of blood transfusion, blood components, or tissue
- 8 = Risk not reported or identified

#### Pediatric exposure categories

- 9 = Hemophilia/coagulation disorder
- 10 = Mother with, or at risk for, HIV infection
- 11 = Receipt of blood transfusion, blood components, or tissue
- 12 = Risk not reported or identified

#### Multirisk (column 20)

*Multirisk* is coded only for adult/adolescent patients (13 years old or older) and indicates if the patient has risk(s) of exposure to HIV other than the one indicated by *exposure*.

- 0 = Patient's only mode of exposure to HIV is that indicated by *exposure*
- 1 = Patient has additional risk(s) of exposure
- 2 = Patient's mode of exposure is not reported or identified

#### Birth (column 21)

- 1 = Patient was born in the United States or its dependencies and possessions, or place of birth was not specified
- 2 = Patient was born outside the United States

#### Heterosexual risk information (columns 22 through 25)

These variables (*sexbi*, *sexiv*, *sexother*, and *sexhiv*) contain additional exposure information for patients infected heterosexually. All 4 variables are coded as follows:

- 0 = no
- 1 = yes
- 9 = missing/unknown

The variable *sexbi* is coded only for women (for men, the variable contains a blank). All 4 variables contain “9” (missing/unknown) for patients with hemophilia, regardless of whether the exposure information is in

fact unknown. This restriction is necessary in order to comply with the Assurance of Confidentiality on page 5. Of the 4,596 AIDS cases reported through December 1995 among adults/adolescents with hemophilia, less than 4 percent also reported heterosexual contact with a person at increased risk for AIDS or HIV infection.

### Adjwgt (columns 26 through 31)

This variable contains an adjustment weight which, when used as a weighting variable in a frequency tabulation, produces tabulations of AIDS cases that are adjusted for delays in case reporting (see page 11 for a discussion of delays in reporting). The weights are based on estimated reporting delay distributions that take into account exposure, geographic, and demographic variations in case reporting. The adjustment weights and the resulting tabulations are not reliable for cases diagnosed during the most recent 6 months. The *Tools* menu contains an adjusted weight option. If you select this option, all subsequent tabulations you request will be weighted accordingly.

### MSA (columns 32 through 35)

Metropolitan area of residence at diagnosis of AIDS is identified for adult/adolescent patients residing in MSAs with 500,000 or more population, according to the latest available official U.S. Bureau of Census estimates. Each MSA is identified by a 4-digit code listed in Appendix B. For adult/adolescent patients residing in an MSA with less than 500,000 population, in a non-metropolitan area, or whose metropolitan area of residence is unknown, and for all pediatric patients, region of residence is identified. The regional codes are:

- 1 = Northeast: Connecticut, Maine, Massachusetts, New Hampshire, New Jersey, New York, Pennsylvania, Rhode Island, and Vermont
- 2 = Midwest: Indiana, Illinois, Iowa, Kansas, Michigan, Minnesota, Missouri, Nebraska, North Dakota, Ohio, South Dakota, and Wisconsin
- 3 = South: Alabama, Arkansas, Delaware, District of Columbia, Florida, Georgia, Kentucky, Louisiana, Maryland, Mississippi, North Carolina, Oklahoma, South Carolina, Tennessee, Texas, Virginia, and West Virginia
- 4 = West: Alaska, Arizona, California, Colorado, Idaho, Hawaii, Montana, Nevada, New Mexico, Oregon, Utah, Washington, and Wyoming
- 5 = U.S. dependencies, possessions, and independent nations in free association with the United States: Guam, Puerto Rico, the U.S. Virgin Islands, and the U.S. Pacific Islands listed on page 8.



# State, MSA, and County Tables

In addition to the rectangular data file discussed in section 2, the *AIDS Public Information Data Set* contains tabular data by state, metropolitan area, and county or health district. These tables consist of frequency tables and 2-way cross tabulations of 8 variables extracted from CDC's national AIDS surveillance data set. For counties or health districts, the data set contains only 1-way tables of 3 variables (*age*, *race/ethnicity*, and *sex*). The data set contains one set of tables for the entire United States, one set for each state and for the District of Columbia, one set for each MSA, and one set for each county or health district. All MSAs with 500,00 or more population are included in the data set. Selected MSAs with populations between 100,000 and 500,000, and selected counties or health districts are included in the data set, based on the data release policies of the individual states.

Data from MSAs with populations between 100,000 and 500,000 are included from Arkansas, Colorado, Connecticut, Delaware, Florida, Georgia, Hawaii, Idaho, Iowa, Illinois, Indiana, Kansas, Kentucky, Louisiana, Maryland, Maine, Michigan, Minnesota, Mississippi, Missouri, Montana, Nebraska, Nevada, New Hampshire, New Jersey, New Mexico, North Carolina, Ohio, Oklahoma, Oregon, Pennsylvania, Rhode Island, South Carolina, Tennessee, Texas, Virginia, Washington, West Virginia, and Wyoming.

Data from individual counties are included from Arkansas, Delaware, Georgia, Hawaii, Indiana, Louisiana, Minnesota, Missouri, Nevada, New Hampshire, New Jersey, Ohio, Oklahoma, Oregon, Pennsylvania, Rhode Island, and Washington.

Data from health districts are included from Florida, Idaho, Kentucky, Mississippi, Montana, Nebraska, New Mexico, South Carolina, Tennessee, Virginia, and West Virginia.

See Appendix B for a list of MSAs. See Appendix C for a list of counties which comprise each health district.

The variables included in the state and MSA tables are:

Variable	Description
<i>age</i>	Age group at diagnosis of the first AIDS-indicator condition
<i>categ</i>	Indicates which of the CDC AIDS case revisions the patient meets
<i>dth_hyr</i>	Half-year of death for patients reported dead
<i>dx_hyr</i>	Half-year of diagnosis of first AIDS-indicator condition
<i>ent_hyr</i>	Half-year in which CDC first received information about the case
<i>exposure</i>	Mode of exposure to HIV
<i>race/ethnicity</i>	Race/ethnicity of patient
<i>sex</i>	Sex of patient

For counties and health districts, 1-way tables are available for the variables *age*, *race/ethnicity*, and *sex*. The values used for the variables are printed below.

## Age

This variable contains the patient's age when he or she was first diagnosed with an AIDS-indicator condition. Ages printed in the documentation file are grouped as follows:

0 - 1  
1 - 12  
13 - 19  
20 - 29  
30 - 39  
40 - 49  
50 - 59  
60 +

## Categ

This variable reflects revisions made to the CDC surveillance definition for AIDS. Only cases meeting the current (1993) surveillance definition are included in the data set. *Categ* indicates whether the patient also meets the pre-1985, 1985, or 1987 surveillance definition, and whether the diagnosis, if it meets the 1987 or 1993 definition, was definitive or presumptive. Cases that meet more than one of these surveillance definitions are classified into the definition category listed first. For more information about the 1993 definition, see *Morbidity and Mortality Weekly Report, Recommendations and Reports*, December 18, 1992.

- 1 = Case meets the pre-1985 surveillance definition
- 2 = Case meets the 1985 surveillance definition
- 3 = Case meets the 1987 surveillance definition and was diagnosed definitively
- 4 = Case meets the 1987 surveillance definition and was diagnosed presumptively
- 5 = Case meets the 1993 surveillance definition: pulmonary tuberculosis, recurrent pneumonia, and/or cervical cancer (definitive diagnosis)
- 6 = Case meets the 1993 surveillance definition: pulmonary tuberculosis and/or recurrent pneumonia (presumptive diagnosis)
- 7 = Cases meets the 1993 surveillance definition: severe HIV-related immunosuppression

## Dth\_hyr

For patients whose death has been reported to CDC, this variable contains the half-year of death. The first four numbers indicate the year; the last two indicate the first or second half of that year. For example, the value "198802" indicates that the patient died in the second half of 1988. Patients whose death has been reported to CDC, but whose date of death is unknown are coded as "999999".

## Dx\_hyr

This variable contains the half-year in which the first AIDS-indicator condition was diagnosed. The first four numbers indicate the year; the last two indicate the first or second half of that year.

## Ent\_hyr

This variable contains the half-year in which CDC received the case report. The first four numbers indicate the year; the last two indicate the first or second half of that year.

## Exposure

For surveillance purposes, AIDS cases are counted only once in a hierarchy of exposure categories. Persons with more than one reported mode of exposure to HIV are classified in the exposure category listed first in the hierarchy, except for men with both a history of sexual contact with other men and injecting drug use. They make up a separate exposure category.

"Men who have sex with men" cases include men who report sexual contact with other men (i.e., homosexual contact) and men who report sexual contact with both men and women (i.e., bisexual contact). "Heterosexual contact" cases are in persons who report specific heterosexual contact with a person with, or at increased risk for, HIV infection (e.g., an injecting drug user).

Adults/adolescents born, or who had sex with someone born, in a country where heterosexual transmission was believed to be the predominant mode of HIV transmission (formerly classified as Pattern-II countries by the World Health Organization) are no longer classified as having heterosexually acquired AIDS. Similar to case reports for other persons who are reported without behavioral or transfusion risks for HIV, these reports are now classified (in the absence of other risk information which would classify them into another exposure category) as "no risk reported or identified" (see *Morbidity and Mortality Weekly Report*,



March 11, 1994). Children whose mother was born, or whose mother had sex with someone born, in a Pattern-II country are now classified (in the absence of other risk information which would classify them into another exposure category) as “Mother with/at risk for HIV infection: has HIV infection, risk not specified.”

“Risk not reported or identified” cases are in persons with no reported history of exposure to HIV through any of the routes listed in the hierarchy of exposure categories. Risk not reported or identified cases include persons who are currently under investigation by local health department officials; persons whose exposure history is incomplete because they died, declined to be interviewed, or were lost to follow-up; and persons who were interviewed or for whom other follow-up information was available and no exposure mode was identified. Persons who have an exposure mode identified at the time of follow-up are reclassified into the appropriate exposure category.

- 01 = Men who have sex with men
- 02 = Injecting drug use
- 03 = Men who have sex with men and inject drugs
- 04 = Adult/adolescent hemophilia/coagulation disorder
- 05 = Heterosexual contact with a person with, or at increased risk for, HIV infection
- 07 = Adult/adolescent receipt of blood transfusion, blood components, or tissue
- 08 = Adult/adolescent risk not reported or identified
- 09 = Pediatric hemophilia/coagulation disorder
- 10 = Mother with, or at risk for, HIV infection
- 11 = Pediatric receipt of blood transfusion, blood components, or tissue
- 12 = Pediatric risk not reported or identified

## Race/ethnicity

- 1 = White (not Hispanic)
- 2 = Black (not Hispanic)
- 3 = Hispanic
- 4 = Asian/Pacific Islander
- 5 = American Indian/Alaskan Native
- 9 = Unknown

## Sex

- 1 = Male
- 2 = Female

## Small Cell Restriction

In accordance with CDC guidelines on protecting confidentiality and with an agreement made with state and local health departments for release of these data, entries whose value is 3 or less are not included in the tables. In addition, the *AIDS Public Information Data Set* software allows you to combine data from more than one state, MSA, or county/health district in either separate or aggregate form. If you select the aggregate option, each count may be off by an amount equal to 3 times the number of states/MSAs/counties aggregated. For example, if you select data from California, Washington State, and Oregon, each count may be off by as many as 9 cases (3 times the number of states, in this case 3).



# Appendix A: Installation

In 1997, The *AIDS Public Information Data Set* was rewritten to be fully Windows compatible. While much of the original program design remained unchanged, many features were added, cursor and mouse controls were enhanced, and the installation procedure changed to reflect Windows conventions. Changes to the software are more fully described in the on-line help screens. The December 1999 edition contains the changes made in 1997. As with previous releases, the software allows you to display simple statistics without additional software such as SAS, SPSS, BMDP, or PRODAS. More complex analyses, however, require statistical software.

To transfer the data to another software package for analysis, you may wish to download only the ASCII version. You may also load the software and use the export option (under *File*) to extract the records and variables you wish to analyze. The export option will create an ASCII data file, which can then be processed by other software.

## Loading the Software

The *AIDS Public Information Data Set* is available on CD, as part of the *CDC HIV/AIDS Information Guide*, or can be downloaded from CDC's World Wide Web. Installation instructions vary, depending on the medium you are using.

Minimum requirements for installation are:

- \* Windows 95 or greater
- \* 80486 CPU
- \* 420 K of free RAM
- \* 50Mb of free disk space

To install the software from CDC's web site

- 1) Download the self-extracting file (PidsInst.exe) to desired directory (i.e., C:\AIDSPIDS).
- 2) Click on Start and Run. Using the Browse feature, locate and run PIDSInst.EXE.

You may change the drive and directory to which the *AIDS Public Information Data Set* will be extracted.

If you want to be able to run the program from the Start Menu, be sure "Create program group(s): AIDS Public Information Data Set" is checked.

- 3) Click on Extract.
- 4) After the program has been extracted, double click on the AIDS Public Information Data Set icon to run it. The first time you run it, it will perform a setup/indexing process that will take up to a few minutes to complete.
- 5) In order to save disk space, the file PIDSINST.exe can be deleted.

To load the software from the CD, insert the disk into the reader. The software will automatically display the initial screen for the *CDC HIV/AIDS Information Guide*. To access the *AIDS Public Information Data Set*, first select menu item 7, "Software." Then select *AIDS Public Information Data Set*. Finally, select "Download PidsInst.exe." This selection will initiate the software installation procedure described above. Simply proceed with steps 1 through 5, above.

## Getting Help

The *AIDS Public Information Data Set* uses standard Windows interfaces, and can be mastered with minimum effort. On-line help screens describe how to use the program to display information. You can access help by pressing the <F1> key, by clicking the right (secondary) mouse button, or by selecting the *Help* menu. The information displayed will vary depending upon the last option you accessed. If you need additional information, contact the Statistics and Data Management Branch, Division of HIV/AIDS Prevention, telephone (404) 639-2020.

## Displaying the Menus

Once you complete the installation procedure and run the program, you will see a screen with four options displayed on the upper-left corner: *File*, *Tools*, *Window*, and *Help*. Select *File* to display data from either the main data file or from the state, MSA, or county tables. A second screen will display so that you may select the variables you wish to tabulate. Select *Tools* to create indexes or set various options that control the display of data. Select *Window* to scroll through the tables you have created. Select *Help* to see further information on how to use this program.

## Cursor Control

Cursor control uses a standard Windows interface. Select variables by double-clicking the left (primary) mouse button or the *Enter* key. An asterisk will display next to the fields you have selected. Menus and

other options can also be selected by pressing the <Alt> key and typing the highlighted letter of that option.

## Appendix B: Metropolitan Statistical Areas

Definitions for MSAs are issued by the Office of Management and Budget (OMB) to be used in presentation of statistics by agencies of the federal government. The metropolitan areas used on the *AIDS Public Information Data Set* are the MSAs for all areas except the 6 New England states. For these states, the New England County Metropolitan Areas (NECMA, also defined by OMB) are used. Metropolitan areas are named for a central city in the MSA or NECMA and may include several counties and cross state boundaries.

The AIDS Public Information Data Set contains data from all MSAs with 500,000 or more population, and from MSAs with 100,000 to 500,000 population from Arkansas, Colorado, Connecticut, Delaware, Florida, Georgia, Hawaii, Idaho, Iowa, Illinois, Indiana, Kansas, Kentucky, Louisiana, Maryland, Maine, Michigan, Minnesota, Mississippi, Missouri, Montana, Nebraska, Nevada, New Hampshire, New Jersey, New Mexico, North Carolina, Ohio, Oklahoma, Oregon, Pennsylvania, Rhode Island, South Carolina, Tennessee, Texas, Virginia, Washington, West Virginia, and Wyoming.

Code	Metropolitan areas with 500,000 or more population
------	--

80	Akron, Ohio
160	Albany-Schenectady, N.Y.
200	Albuquerque, N.Mex.
240	Allentown, Pa.
440	Ann Arbor, Mich.
520	Atlanta, Ga.
640	Austin, Tex.
680	Bakersfield, Calif.
720	Baltimore, Md.
760	Baton Rouge, La.
875	Bergen-Passaic, N.J.
1000	Birmingham, Ala.
1123	Boston, Mass.
1280	Buffalo, N.Y.
1520	Charlotte, N.C.
1600	Chicago, Ill.
1640	Cincinnati, Ohio
1680	Cleveland, Ohio
1840	Columbus, Ohio
1920	Dallas, Tex.

2000	Dayton, Ohio
2080	Denver, Colo.
2160	Detroit, Mich.
Code	Metropolitan areas with 500,000 or more population
2320	El Paso, Tex.
2680	Fort Lauderdale, Fla.
2800	Fort Worth, Tex.
2840	Fresno, Calif.
2960	Gary, Ind.
3000	Grand Rapids, Mich.
3120	Greensboro, N.C.
3160	Greenville, S.C.
3240	Harrisburg, Pa.
3283	Hartford, Conn.
3320	Honolulu, Hawaii
3360	Houston, Tex.
3480	Indianapolis, Ind.
3600	Jacksonville, Fla.
3640	Jersey City, N.J.
3760	Kansas City, Mo.
3840	Knoxville, Tenn.
4120	Las Vegas, Nev.
4400	Little Rock, Ark.
4480	Los Angeles, Calif.
4520	Louisville, Ky.
4880	McAllen, Tex.
4920	Memphis, Tenn.
5000	Miami, Fla.
5015	Middlesex, N.J.
5080	Milwaukee, Wis.
5120	Minneapolis-Saint Paul, Minn.
5160	Mobile, Ala.
5190	Monmouth-Ocean City, N.J.
5360	Nashville, Tenn.
5380	Nassau-Suffolk, N.Y.
5483	New Haven, Conn.
5560	New Orleans, La.
5600	New York, N.Y.
5640	Newark, N.J.

5720	Norfolk, Va.
5775	Oakland, Calif.
5880	Oklahoma City, Okla.
5920	Omaha, Nebr.
5945	Orange County, Calif.
5960	Orlando, Fla.
Code	Metropolitan areas with 500,000 or more population
6160	Philadelphia, Pa.
6200	Phoenix, Ariz.
6280	Pittsburgh, Pa.
6440	Portland, Oreg.
6483	Providence, R.I.
6640	Raleigh-Durham, N.C.
6760	Richmond, Va.
6780	Riverside-San Bernardino, Calif.
6840	Rochester, N.Y.
6920	Sacramento, Calif.
7040	Saint Louis, Mo.
7160	Salt Lake City, Utah
7240	San Antonio, Tex.
7320	San Diego, Calif.
7360	San Francisco, Calif.
7400	San Jose, Calif.
7440	San Juan, P.R.
7510	Sarasota, Fla.
7560	Scranton, Pa.
7600	Seattle, Wash.
8003	Springfield, Mass.
8120	Stockton, Calif.
8160	Syracuse, N.Y.
8200	Tacoma, Wash.
8280	Tampa-Saint Petersburg, Fla.
8400	Toledo, Ohio
8520	Tucson, Ariz.
8560	Tulsa, Okla.
8735	Ventura, Calif.
8840	Washington, D.C.
8960	West Palm Beach, Fla.
9040	Wichita, Kans.



9160	Wilmington, Del.
9243	Worcester, Mass.
9320	Youngstown, Ohio

Code	Metropolitan areas with 100,000 to 500,000 population
------	---

40	Abilene, Tex.
120	Albany, Ga.

Code	Metropolitan areas with 100,000 to 500,000 population
------	---

220	Alexandria, La.
280	Altoona, Pa.
320	Amarillo, Tex.
480	Asheville, N.C.
500	Athens, Ga.
560	Atlantic-Cape May, N.J.
600	Augusta, Ga.
733	Bangor, Maine
840	Beaumont, Tex.
860	Bellingham, Wash.
870	Benton Harbor, Mich
880	Billings, Mont.
920	Biloxi, Miss.
1020	Bloomington, Ind.
1040	Bloomington, Ill.
1080	Boise, Idaho
1125	Boulder, Colo.
1145	Brazoria, Tex.
1150	Bremerton, Wash.
1240	Brownsville, Tex.
1260	Bryan, Tex.
1320	Canton, Ohio
1360	Cedar Rapids, Iowa
1400	Champaign-Urbana, Ill.
1440	Charleston, S.C.
1480	Charleston, W. Va.
1540	Charlottesville, Va.
1560	Chattanooga, Tenn.
1660	Clarksville, Tenn.

1720	Colorado Springs, Colo.
1740	Columbia, Mo.
1760	Columbia, S.C.
1800	Columbus, Ga.
1880	Corpus Christi, Tex.
1900	Cumberland, Md.
1950	Danville, Va.
1960	Davenport, Iowa
2020	Daytona Beach, Fla.
2040	Decatur, Ill.
2120	Des Moines, Iowa
2190	Dover, Del.

Code	Metropolitan areas with 100,000 to 500,000 population
------	---

2240	Duluth, Minn.
2330	Elkhart, Ind.
2360	Erie, Penn.
2400	Eugene, Oreg.
2440	Evansville, Ind.
2520	Fargo, N.D.
2560	Fayetteville, N.C.
2580	Fayetteville, Ark.
2640	Flint, Mich.
2655	Florence, S.C.
2670	Fort Collins, Colo.
2700	Fort Myers, Fla.
2710	Fort Pierce, Fla.
2720	Fort Smith, Ark.
2750	Fort Walton Beach, Fla.
2760	Fort Wayne, Ind.
2900	Gainesville, Fla.
2920	Galveston, Tex.
2980	Goldsboro, N.C.
2995	Grand Junction, Colo.
3060	Greeley, Colo.
3150	Greenville, N.C.
3180	Hagerstown, Md.
3200	Hamilton, Ohio
3285	Hattiesburg, Miss.
3290	Hickory, N.C.

3350	Houma, La.
3400	Huntington, W.Va.
3500	Iowa City, Iowa
3520	Jackson, Mich.
3560	Jackson, Miss.
3580	Jackson, Tenn.
3605	Jacksonville, N.C.
3660	Johnson City, Tenn.
3680	Johnstown, Pa.
3710	Joplin, Mo.
3720	Kalamozoo, Mich.
3740	Kankakee, Ill.
3810	Killeen, Tex.
3850	Kokomo, Ind.
3880	Lafayette, La.

Code	Metropolitan areas with 100,000 to 500,000 population
------	---

3920	Lafayette, Ind.
3960	Lake Charles, La.
3980	Lakeland, Fla.
4000	Lancaster, Pa.
4040	Lansing, Mich.
4080	Laredo, Tex.
4100	Las Cruces, N.Mex.
4200	Lawton, Okla.
4243	Lewiston, Maine
4280	Lexington, Ky.
4320	Lima, Ohio
4360	Lincoln, Nebr.
4420	Longview, Tex.
4600	Lubbock, Tex.
4640	Lynchburg, Va.
4680	Macon, Ga.
4800	Mansfield, Ohio
4890	Medford, Oreg.
4900	Melbourne, Fla.
5200	Monroe, La.
5280	Muncie, Ind.
5330	Myrtle Beach, S.C.
5345	Naples, Fla.

5523	New London, Conn.
5660	Newburgh, N.Y.
5790	Ocala, Fla.
5800	Odessa, Tex.
5910	Olympia, Wash.
6015	Panama City, Fla.
6020	Parkersburg, W.Va.
6080	Pensacola, Fla.
6120	Peoria, Ill.
6403	Portland, Maine
6560	Pueblo, Colo.
6580	Punta Gorda, Fla.
6680	Reading, Pa.
6720	Reno, Nev.
6740	Richland, Wash.
6800	Roanoka, Va.
6820	Rochester, Minn.

Code	Metropolitan areas with 100,000 to 500,000 population
------	---

6880	Rockford, Ill.
6895	Rocky Mount, N.C.
6960	Saginaw, Mich.
6980	Saint Cloud, Minn.
7080	Salem, Oreg.
7200	San Angelo, Tex.
7490	Santa Fe, N.Mex.
7520	Savannah, Ga.
7610	Sharon, Pa.
7640	Sherman, Tex.
7680	Shreveport, La.
7720	Sioux City, Iowa
7800	South Bend, Ind.
7840	Spokane, Wash.
7880	Springfield, Ill.
7920	Springfield, Mo.
8050	State College, Pa.
8080	Steubenville, Ohio
8140	Sumter, S.C.
8240	Tallahassee, Fla.
8320	Terre Haute, Ind.

8360	Texarkana, Tex.
8440	Topeka, Kans.
8480	Trenton, N.J.
8640	Tyler, Tex.
8720	Vallejo, Calif.
8760	Vineland, N.J.
8800	Waco, Tex.
8920	Waterloo, Iowa
9000	Wheeling, W.Va.
9080	Wichita Falls, Tex.
9140	Williamsport, Pa.
9200	Wilmington, N.C.
9260	Yakima, Wash.
9280	York, Pa.



# Appendix C: Health Districts

Listed below are the counties which comprise each health district included in the data set. The county name is preceded by its Federal Information Processing Standards (FIPS) code (see *Worldwide Geographic Location Codes*, available from the General Services Administration, telephone 202-219-0077).

## Florida

### Health District 1

Health District 1		67	Lafayette
		75	Levy
33	Escambia	107	Putnam
91	Okaloosa	121	Suwannee
113	Santa Rosa	125	Union
131	Walton		

### Health District 2

		3	Baker
5	Bay	19	Clay
13	Calhoun	31	Duval
37	Franklin	89	Nassau
39	Gadsden	109	Saint Johns
45	Gulf		
59	Holmes		
63	Jackson		Health District 5
65	Jefferson		
73	Leon	101	Pasco
77	Liberty	103	Pinellas
79	Madison		
123	Taylor		Health District 6
129	Wakulla		
133	Washington	57	Hillsborough

### Health District 3

1	Alachua		
7	Bradford		
23	Columbia		
29	Dixie		
41	Gilchrist		
47	Hamilton		

### Health District 4

	3	Baker
	19	Clay
	31	Duval
	89	Nassau
	109	Saint Johns

### Health District 5

101	Pasco
	Pinellas

### Health District 6

57	Hillsborough
81	Manatee

### Health District 7

9	Brevard
95	Orange
97	Osceola
117	Seminole

### Health District 8

15	Charlotte
21	Collier
27	De Soto
43	Glades
51	Hendry
71	Lee
115	Sarasota

### Health District 9

99	Palm Beach
----	------------

### Health District 10

11	Broward
----	---------

### Health District 11

25	Dade
87	Monroe

### Health District 12

35	Flagler
127	Volusia

# Florida

## Health District 13

17	Citrus
53	Hernando
69	Lake
83	Marion
119	Sumter

## Health District 14

49	Hardee
55	Highlands
105	Polk

## Health District 15

61	Indian River
85	Martin
93	Okeechobee
111	Saint Lucie



# Idaho

## Health District 1

9 Benewah  
17 Bonner  
21 Boundary  
55 Kootenai  
79 Shoshone

## Health District 2

35 Clearwater  
49 Idaho  
57 Latah  
61 Lewis  
69 Nez Perce

## Health District 3

3 Adams  
27 Canyon  
45 Gem  
73 Owyhee  
75 Payette  
87 Washington

## Health District 4

1 Ada  
15 Boise  
39 Elmore  
85 Valley

## Health District 5

13 Blaine  
25 Camas  
31 Cassia  
47 Gooding  
53 Jerome  
63 Lincoln  
67 Minidoka  
83 Twin Falls

## Health District 6

5 Bannock  
7 Bear Lake  
11 Bingham  
23 Butte  
29 Caribou  
41 Franklin  
71 Oneida  
77 Power

## Health District 7

19 Bonneville  
33 Clark  
37 Custer  
43 Fremont  
51 Jefferson  
59 Lemhi  
65 Madison  
81 Teton

# Kentucky

## Health District 1

7 Ballard  
35 Calloway  
39 Carlisle  
75 Fulton  
83 Graves  
105 Hickman  
145 McCracken  
157 Marshall

## Health District 2

33 Caldwell  
47 Christian  
55 Crittenden  
107 Hopkins  
139 Livingston  
143 Lyon  
177 Muhlenberg  
219 Todd  
221 Trigg

## Health District 3

59 Daviess  
91 Hancock  
101 Henderson  
149 McLean  
183 Ohio  
225 Union  
233 Webster

## Health District 4

3 Allen  
9 Barren  
31 Butler  
61 Edmonson

99 Hart  
141 Logan  
169 Metcalfe  
171 Monroe  
213 Simpson  
227 Warren

## Health District 5

27 Breckinridge  
85 Grayson  
93 Hardin  
123 Larue  
155 Marion  
163 Meade  
179 Nelson  
229 Washington

## Health District 6

29 Bullitt  
103 Henry  
111 Jefferson  
185 Oldham  
211 Shelby  
215 Spencer  
223 Trimble

## Health District 7

15 Boone  
37 Campbell  
41 Carroll  
77 Gallatin  
81 Grant  
117 Kenton  
187 Owen  
191 Pendleton

## Health District 8

23 Bracken  
69 Fleming  
135 Lewis  
161 Mason  
201 Robertson

## Health District 9

11 Bath  
165 Menifee  
173 Montgomery  
175 Morgan  
205 Rowan

## Health District 10

19 Boyd  
43 Carter  
63 Elliott  
89 Greenup  
127 Lawrence

## Health District 11

71 Floyd  
115 Johnson  
153 Magoffin  
159 Martin  
195 Pike

# Kentucky

## Health District 12

25 Breathitt  
119 Knott  
129 Lee  
131 Leslie  
133 Letcher  
189 Owsley  
193 Perry  
237 Wolfe

## Health District 13

13 Bell  
51 Clay  
95 Harlan  
109 Jackson  
121 Knox  
125 Laurel  
203 Rockcastle  
235 Whitley

## Health District 14

1 Adair  
45 Casey  
53 Clinton  
57 Cumberland  
87 Green  
147 McCreary  
199 Pulaski  
207 Russell  
217 Taylor  
231 Wayne

## Health District 15

5 Anderson  
17 Bourbon  
21 Boyle  
49 Clark  
65 Estill  
67 Fayette  
73 Franklin  
79 Garrard  
97 Harrison  
113 Jessamine  
137 Lincoln  
151 Madison  
167 Mercer  
181 Nicholas  
197 Powell  
209 Scott  
239 Woodford

# Mississippi

## Health District 1

27 Coahoma  
33 De Soto  
43 Grenada  
107 Panola  
119 Quitman  
135 Tallahatchie  
137 Tate  
143 Tunica  
161 Yalobusha

## Health District 4

13 Calhoun  
17 Chickasaw  
19 Choctaw  
25 Clay  
87 Loundes  
95 Monroe  
103 Noxubee  
105 Oktibbeha  
155 Webster  
159 Winston

## Health District 7

1 Adams  
5 Amite  
37 Franklin  
63 Jefferson  
77 Lawrence  
85 Lincoln  
113 Pike  
147 Walthall  
157 Wilkinson

## Health District 2

3 Alcorn  
9 Benton  
57 Itawamba  
71 Lafayette  
81 Lee  
93 Marshall  
115 Pontotoc  
117 Prentiss  
139 Tippah  
141 Tishomingo  
145 Union

## Health District 5

21 Claiborne  
29 Copiah  
49 Hinds  
55 Issaquena  
89 Madison  
121 Rankin  
125 Sharkey  
127 Simpson  
149 Warren  
163 Yazoo

## Health District 8

31 Covington  
35 Forrest  
41 Greene  
65 Jefferson Davis  
67 Jones  
73 Lamar  
91 Marion  
111 Perry  
153 Wayne

## Health District 3

7 Attala  
11 Bolivar  
15 Carroll  
51 Holmes  
53 Humphreys  
83 Leflore  
97 Montgomery  
133 Sunflower  
151 Washington

## Health District 6

61 23 Clarke  
Jasper  
69 Kemper  
75 Lauderdale  
79 Leake  
99 Neshoba  
101 Newton  
123 Scott  
129 Smith

## Health District 9

39 George  
45 Hancock  
47 Harrison  
59 Jackson  
109 Pearl River  
131 Stone

# Montana

## Health District 1

11 Carter  
17 Custer  
19 Daniels  
21 Dawson  
25 Fallon  
33 Garfield  
55 McCone  
71 Phillips  
75 Powder River  
79 Prairie  
83 Richland  
85 Roosevelt  
87 Rosebud  
91 Sheridan  
103 Treasure  
105 Valley  
109 Wibaux

## Health District 2

5 Blaine  
13 Cascade  
15 Chouteau  
35 Glacier  
41 Hill  
51 Liberty  
73 Pondera  
99 Teton  
101 Toole

## Health District 3

3 Big Horn  
9 Carbon  
27 Fergus  
37 Golden Valley  
45 Judith Basin  
65 Musselshell  
69 Petroleum  
95 Stillwater  
97 Sweet Grass  
107 Wheatland  
111 Yellowstone

## Health District 4

1 Beaverhead  
7 Broadwater  
23 Deer Lodge  
31 Gallatin  
39 Granite  
43 Jefferson  
49 Lewis and Clark  
57 Madison  
59 Meagher  
67 Park  
77 Powell  
93 Silver Bow

## Health District 5

29 Flathead  
47 Lake  
53 Lincoln

## Health District 6

61 Mineral  
63 Missoula  
81 Ravalli  
89 Sanders

# Nebraska

## Health District 1

3 Antelope  
11 Boone  
15 Boyd  
17 Brown  
21 Burt  
27 Cedar  
31 Cherry  
37 Colfac  
39 Cuming  
43 Dakota  
51 Dixon  
89 Holt  
103 Keya Paha  
107 Knox  
119 Madison  
125 Nance  
139 Pierce  
141 Platte  
149 Rock  
167 Stanton  
173 Thurston  
179 Wayne

## Health District 2

53 Dodge  
55 Douglas  
153 Sarpy  
177 Washington

## Health District 3

23 Butler  
25 Cass  
59 Fillmore  
67 Gage  
95 Jefferson  
97 Johnson

109 Lancaster  
125 Nemaha  
131 Otoe  
133 Pawnee  
143 Polk  
147 Richardson  
151 Saline  
155 Saunders  
159 Seward  
169 Thayer  
185 York

## Health District 4

1 Adams  
9 Blaine  
19 Buffalo  
35 Clay  
41 Custer  
61 Franklin  
71 Garfield  
77 Greeley  
79 Hall  
81 Hamilton  
83 Harlan  
93 Howard  
99 Kearney  
115 Loup  
121 Merrick  
129 Nuckolls  
137 Phelps  
163 Sherman  
175 Valley  
181 Webster  
183 Wheeler

## Health District 5

5 Arther  
29 Chase  
47 Dawson  
57 Dundy  
63 Frontier  
65 Furnas  
73 Gosper  
75 Grant  
85 Hayes  
87 Hitchcock  
91 Hooker  
101 Keith  
111 Lincoln  
113 Logan  
117 McPherson  
135 Perkins  
145 Red Willow

## Health District 6

7 Banner  
13 Box Butte  
33 Cheyenne  
45 Dawes  
49 Deuel  
69 Garden  
105 Kimball  
123 Morrill  
137 Scotts Bluff  
161 Sheridan  
165 Sioux

# New Mexico

## Health District 1

1 Bernalillo  
6 Cibola  
31 McKinley  
43 Sandoval  
45 San Juan  
57 Torrance  
61 Valencia

## Health District 2

7 Colfax  
21 Harding  
28 Los Alamos  
33 Mora  
39 Rio Arriba  
47 San Miguel  
49 Santa Fe  
55 Taos  
59 Union

## Health District 3

3 Catron  
13 Dona Ana  
17 Grant  
23 Hidalgo  
29 Luna  
35 Otero  
51 Sierra  
53 Socorro

## Health District 4

5 Chaves  
9 Curry  
11 De Baca  
15 Eddy  
19 Guadalupe  
25 Lea  
27 Lincoln  
37 Quay  
41 Roosevelt

# South Carolina

## Health District 1

7 Anderson  
73 Oconee

## Health District 2

45 Greenville  
77 Pickens

## Health District 3

21 Cherokee  
83 Spartanburg  
87 Union

## Health District 4

23 Chester  
57 Lancaster  
91 York

## Health District 5

9 Bamberg  
17 Calhoun  
75 Orangeburg

## Health District 6

13 Beaufort  
29 Colleton  
49 Hampton  
53 Jasper

## Health District 7

3 Aiken  
5 Allendale  
11 Barnwell

## Health District 8

39 Fairfield  
63 Lexington  
71 Newberry  
79 Richland

## Health District 9

25 Chesterfield  
31 Darlington  
69 Marlboro

## Health District 9

33 Dillon  
41 Florence  
67 Marion

## Health District 10

15 Berkeley  
19 Charleston  
35 Dorchester

## Health District 11

1 Abbeville  
37 Edgefield  
47 Greenwood  
59 Laurens  
65 McCormick  
81 Saluda

## Health District 12

43 Georgetown  
51 Horry  
89 Williamsburg

## Health District 13

27 Clarendon  
55 Kershaw  
61 Lee  
85 Sumter



# Tennessee

## Health District 1

5 Benton  
17 Carroll  
23 Chester  
33 Crockett  
39 Decatur  
45 Dyer  
47 Fayette  
53 Gibson  
69 Hardeman  
71 Hardin  
75 Haywood  
77 Henderson  
79 Henry  
95 Lake  
97 Lauderdale  
109 McNairy  
131 Obion  
167 Tipton  
183 Weakley

## Health District 2

21 Cheatham  
43 Dickson  
83 Houston  
85 Humphreys  
125 Montgomery  
147 Robertson  
149 Rutherford  
161 Stewart  
165 Sumner  
169 Trousdale  
187 Williamson  
189 Wilson

## Health District 3

3 Bedford  
31 Coffee  
55 Giles  
81 Hickman  
99 Lawrence  
101 Lewis  
103 Lincoln  
117 Marshall  
119 Maury  
127 Moore  
135 Perry  
181 Wayne

## Health District 4

15 Cannon  
27 Clay  
35 Cumberland  
41 Dekalb  
49 Fentress  
87 Jackson  
111 Macon  
133 Overton  
137 Pickett  
141 Putnam  
159 Smith  
175 Van Buren  
177 Warren  
185 White

## Health District 5

7 Bledsoe  
11 Bradley  
51 Franklin  
61 Grundy  
107 McMinn  
115 Marion  
121 Meigs  
139 Polk  
143 Rhea  
153 Sequatchie

## Health District 6

1 Anderson  
9 Blount  
13 Campbell  
25 Claiborne  
29 Cocke  
57 Grainger  
59 Greene  
63 Hamblen  
89 Jefferson  
105 Loudon  
123 Monroe  
129 Morgan  
145 Roane  
151 Scott  
155 Seiver  
173 Union

# Tennessee

## Health District 7

19 Carter  
67 Hancock  
73 Hawkins  
91 Johnson  
171 Unicoi  
179 Washington

## Health District 8

157 Shelby

## Health District 9

113 Madison

## Health District 10

37 Davidson

## Health District 11

65 Hamilton

## Health District 12

93 Knox

## Health District 13

163 Sullivan

# Virginia

## Health District 1

3 Albemarle  
 15 Augusta  
 17 Bath  
 33 Caroline  
 43 Clarke  
 47 Culpeper  
 61 Fauquier  
 65 Fluvanna  
 69 Frederick  
 79 Greene  
 91 Highland  
 99 King George  
 109 Louisa  
 113 Madison  
 125 Nelson 31  
 137 Orange  
 139 Page  
 157 Rappahannock  
 163 Rockbridge  
 165 Rockingham  
 171 Shenandoah  
 177 Spotsylvania 77  
 179 Stafford  
 187 Warren  
 530 Buena Vista City  
 540 Charlottesville City  
 630 Fredericksburg City  
 660 Harrisonburg City  
 678 Lexington City  
 790 Staunton City 167  
 820 Waynesboro City  
 840 Winchester City

## Health District 2

13 Arlington  
 59 Fairfax  
 107 Loudoun  
 153 Prince William 580

510 Alexandria City  
 600 Fairfax City  
 610 Falls Church City  
 683 Manassas City  
 685 Manassas Park City

## Health District 3

5 Alleghany  
 9 Amherst  
 11 Appomattox  
 19 Bedford  
 21 Bland  
 23 Botetourt  
 27 Buchanan  
 Campbell  
 35 Carroll  
 45 Craig  
 51 Dickenson  
 63 Floyd  
 67 Franklin  
 71 Giles  
 Grayson  
 89 Henry  
 105 Lee  
 121 Montgomery  
 141 Patrick  
 143 Pittsylvania  
 155 Pulaski  
 161 Roanoke  
 Russell  
 169 Scott  
 173 Smyth  
 185 Tazewell  
 191 Washington  
 195 Wise  
 197 Wythe  
 515 Bedford City  
 520 Bristol City  
 560 Clifton Forge City  
 Covington City

590 Danville City  
 640 Galax City  
 680 Lynchburg City  
 690 Martinsville City  
 720 Norton City  
 750 Radford City  
 770 Roanoke City  
 775 Salem City

## Health District 4

7 Amelia  
 25 Brunswick  
 29 Buckingham  
 36 Charles City  
 37 Charlotte  
 41 Chesterfield  
 49 Cumberland  
 53 Dinwiddie  
 75 Goochland  
 81 Greensville  
 83 Halifax  
 85 Hanover  
 87 Henrico  
 111 Lunenburg  
 117 Mecklenburg  
 127 New Kent  
 135 Nottoway  
 145 Powhatan  
 147 Prince Edward  
 Prince George  
 181 Surry  
 183 Sussex  
 570 Colonial Heights City  
 595 Emporia City  
 670 Hopewell City  
 730 Petersburg City  
 760 Richmond City

# Virginia

## Health District 5

1	Accomack
57	Essex
73	Gloucester
93	Isle of Wight
95	James City
97	King and Queen
101	King William
103	Lancaster
115	Mathews
119	Middlesex
131	Northampton
133	Northumberland
159	Richmond
175	Southampton
193	Westmoreland
199	York
550	Chesapeake City
620	Franklin City
650	Hampton City
700	Newport News City
710	Norfolk City
735	Poquoson City
740	Portsmouth City
800	Suffolk City
810	Virginia Beach City
830	Williamsburg City

# West Virginia

## Health District 1

47 McDowell  
55 Mercer  
63 Monroe  
81 Raleigh  
89 Summers  
109 Wyoming

## Health District 2

11 Cabell  
43 Lincoln  
45 Logan  
53 Mason  
59 Mingo  
99 Wayne

## Health District 3

5 Boone  
15 Clay  
39 Kanawha  
79 Putnam

## Health District 4

7 Braxton  
19 Fayette  
25 Greenbrier  
67 Nicholas  
75 Pocahontas  
101 Webster

## Health District 5

13 Calhoun  
35 Jackson  
73 Pleasants  
85 Ritchie  
87 Roane  
95 Tyler  
105 Wirt  
107 Wood

## Health District 6

9 Brooke  
29 Hancock  
51 Marshall  
69 Ohio  
103 Wetzel

## Health District 7

1 Barbour  
17 Doddridge  
21 Gilmer  
33 Harrison  
41 Lewis  
49 Marion  
61 Monongalia  
77 Preston  
83 Randolph  
91 Taylor  
93 Tucker  
97 Upshur

## Health District 8

3 Berkeley  
23 Grant  
27 Hampshire  
31 Hardy  
37 Jefferson  
57 Mineral  
65 Morgan  
71 Pendleton